AMENDMENTS TO THE DRAWINGS

Please replace FIGs. 1A, 1B, 1C, 2A, 2B, 2C, 3, 4A, 4B, and 5 on ten original drawing sheets with FIGs. 1A, 1B, 1C, 2A, 2B, 2C, 3, 4A, 4B, and 5 on ten replacement drawing sheets, which are attached in Appendix I.

Nine annotated drawing sheets showing the changes to FIGs. 1A, 1B, 1C, 2A, 2B, 2C, 4A, 4B, and 5 are provided in Appendix II. FIG. 3 was amended solely to improve the quality of the shading in the bar graph, but otherwise was not changed; therefore, an annotated copy of FIG. 3 is not provided.

REMARKS

Claims 1-18 and 24-28 were pending in this application. Claims 1, 3, 7, 9, 13, and 14 have been amended to correct an obvious clerical error. A substitute specification (excluding the claims) and replacement drawings are submitted to correct the same obvious clerical error identified in the prior-filed specification and drawings. A marked-up version of the specification also is submitted. The marked-up specification shows "all changes relative to the immediately prior version of the specification of record" (37 C.F.R 1.125(c)), which Applicants believe is a version of the specification including the amendments made by Preliminary Amendment, filed July 8, 2003. Accordingly, the July 8, 2003 amendments to the specification are included, but are not shown, in the present mark-up.

Applicants have recently recognized that INF-α21b was mistakenly referred to as INF-α21a throughout the specification and drawings. As was known in the art at the time of filing, INF-α21a and INF-α21b have the same amino acid sequence except for residue 96, which is Met in INF-α21a and Leu in INF-α21b. Applicants attach in Appendix V a copy of Nyman *et al., Biochem. J.*, 329:295-302, January 15, 1998 ("Nyman") to evidence this understanding in the art, and, for instance, refer the examiner to the last two entries of Table 4 in Nyman. As shown, for example, in original FIG. 5 and SEQ ID NO: 34 of the original Sequence Listing, residue 96 of the polypeptide originally identified as "INFα21a" is Leu. Similarly, residue 95 of SEQ ID NO: 13, which corresponds to residue 96 of "INFα21a" by the "facilitated alignment numbering system" (see the specification at page 11, line 32 through page 12, line 11), is Leu. Because residue 96 of IFNα21a would be Met, the "INFα21a" designation is clearly erroneous and the proper designation clearly should be "INFα21b," as now recited throughout the replacement specification, and amended drawings and claims.

No new matter is introduced by the amendments herein.

CONCLUSION

Examination of the pending claims is respectfully requested. The Examiner is invited to call the undersigned if the Examiner believes that a telephone interview would facilitate examination of this application.

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

By _Debra Gordon_

Debra A. Gordon, Ph.D.
Registration No. 54,128

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

APPENDIX I

Replacement Drawing Sheets

[see attached 10 pages]

APPENDIX II

Annotated Drawing Sheets

[see attached 9 pages]

APPENDIX V

Nyman Reference

[see attached 8 pages]

Identification of nine interferon-α subtypes produced by Sendai virus-induced human peripheral blood leucocytes

Tuula A. NYMAN*, Hannele TÖLÖ†, Jaakko PARKKINEN† and Nisse KALKKINEN*[1]

*Institute of Biotechnology, Protein Chemistry Laboratory, P.O. Box 56 (Viikinkaari 9), FIN-00014 University of Helsinki, Finland, and †Finnish Red Cross Blood Transfusion Service, Kivihaantie 7, FIN-00310 Helsinki, Finland

The human interferon-α (IFN-α) family is encoded by 13 different functional genes, and including all cloned sequence variants there are 28 potential IFN-α proteins. To find out which of the described sequences are expressed in normal human leucocytes, we have isolated and partly characterized the components of a highly purified IFN-α preparation produced by Sendai virus-induced human peripheral blood leucocytes. The identification protocol consisted of N-terminal sequencing and mass mapping of the proteins separated by reverse-phase HPLC and/or SDS/-PAGE. The highly purified leucocyte IFN-α preparation was found to contain at least nine different IFN-α species: IFN-α1a, IFN-α2b, IFN-α4b, IFN-α7a, IFN-α8b, IFN-α10a, IFN-α14c, IFN-α17b and IFN-α21b. IFN-α1a was the major subtype, comprising approx. 30% of total leucocyte IFN-α. IFN-α14c, the only subtype containing potential N-glycosylation sites, was shown to be glycosylated at Asn-72. Molecular mass determination of the intact proteins by electrospray ionization MS showed that there are no other post-translational modifications in the IFN-α subtypes than the glycosylation of IFN-α2b and IFN-α14c. Only one sequence variant was found for each subtype, suggesting that the other described gene sequences represent allelic variants or mutations that are more rarely found in the general population.

INTRODUCTION

Human interferon-α (IFN-α) comprises a family of extracellular signalling proteins with antiviral, antiproliferating and immunomodulatory activities. They are produced by peripheral blood leucocytes [1], lymphoblastoid [2] and myeloblastoid cell lines [3] on viral activation. IFN-α proteins are encoded by a multigene family comprising 13 genes clustered on human chromosome 9 [4]. Two of these genes, IFN-A1 and IFN-A13, have identical coding sequences and give rise to a single protein species [5]. The similarity between the other IFN-α subtypes is between 78% and 95% at the protein level, and 79 of the 166 amino acids in the IFN-α family are conserved [6]. Most of the IFN-α genes are expressed at the mRNA level in leucocytes induced by Sendai virus [7]. The presence of several IFN-α subtypes in leucocyte culture supernatant after induction by Sendai virus has been reported [8–11], but only IFN-α2 has previously been characterized in detail [9,12]. The biological significance of the expression of several closely similar IFN-α proteins is not known, but several reports suggest that they show quantitatively distinct patterns of antiviral, growth inhibitory, and killer cell-stimulatory activities [13–18]. Because there are potential differences between different IFN-α subtypes it is of importance to biochemically characterize the subtypes and their possible post-translational modifications.

Among the 13 IFN-α genes, a total of 28 different sequence variants have been described [19]. This increases considerably the number of potential IFN-α proteins. These variants differ from each other in one to four amino acid positions. So far only IFN-α2b and IFN-α2c have been shown to be allelic variants,

IFN-α2b being the predominant allele and IFN-α2c only a minor (less than 0.1%) allelic variant [20]. Many of the IFN-α gene variants have been cloned and sequenced from transformed cell lines and they can present mutations not present in the normal leucocyte genome [20–23]. Two IFN-α variants, IFN-α2a and IFN-α2b, are mass-produced in Escherichia coli by recombinant technology and marketed as drugs. Unlike natural leucocyte IFN-α, these recombinant IFN-α products have proved to be immunogenic in some patients [24,25], which could be due to the presence of unnatural forms of IFN-α proteins. Therefore it might be important for the development of IFN-α drugs to establish which IFN-α subtypes and variants are expressed in normal human leucocytes. In this study we have separated and partly characterized IFN-α proteins from a highly purified IFN-α preparation produced in peripheral blood leucocytes by Sendai virus induction. Nine IFN-α subtypes were identified, and only one variant was found for each of them.

EXPERIMENTAL

Materials

Leucocyte IFN-α was produced in Sendai virus-induced human peripheral blood leucocytes, essentially as described before [1] and purified by immunoadsorption chromatography with two monoclonal antibodies and by gel filtration (H. Tölö, unpublished work). The purified material represents a bulk drug developed by the Finnish Red Cross Blood Transfusion Service, which is aimed at clinical use.

Acrylamide/bisacrylamide, ammonium persulphate, α-cyano-4-hydroxycinnamic acid (α-CCA), 3-(cyclohexylamino)propane-1-sulphonic acid (Caps), dithiothreitol (DTT), N,N,N',N'-tetramethylethylenediamine and Tris/HCl were obtained from Sigma (St. Louis, MO, U.S.A.). SDS was from BDH Laboratory Supplies (Poole, Dorset, U.K.). Ammonium bicarbonate and trifluoroacetic acid (TFA) were from FlukaBioChemica (Buchs, Switzerland). Octyl glucopyranoside (ULTROL7* grade) was purchased from Calbiochem* (La Jolla, CA, U.S.A.). 3,5-Dimethoxy-4-hydroxycinnamic acid (sinapinic acid) was from Aldrich (Steinheim, Germany). Acetonitrile (ACN) was from J. T. Baker (Low Water, NJ, U.S.A.) and methanol was from Merck (Darmstadt, Germany). All reagents for N-terminal sequencing were protein sequencer grade reagents from Applied Biosystems (Perkin Elmer, Foster City, CA, U.S.A.). Coomassie Brilliant Blue stain solutions were prepared from PhastGel*Blue R (Pharmacia Biotech, Uppsala, Sweden). All reagents for capillary isoelectric focusing analyses were CE-grade reagents from Bio-Rad (Hercules, CA, U.S.A.), except that Ampholine*5-8 was from Pharmacia Biotech.

Reverse-phase (RP) chromatography

Interferons (1×10^7 i.u., corresponding to 90 µg of protein) were separated by RP-HPLC with a 2.0×150 mm Delta-Pak* HPI C_4 column (5 µm, 30 nm; Waters, Millford, MA, U.S.A.) and a linear gradient of 40-55% solvent B in 150 min. Solvent A was 0.1% TFA and solvent B was 0.075% TFA in ACN. The flow rate was 150 µl/min and UV detection was at 214 nm.

Peptides from interferon digests were separated on a 1.0 mm × 150 mm Vydac C_8 column (5 µm, 30 nm; LC-Packings, Amsterdam, The Netherlands) with the same solvents as above. Elution was done with linear gradients of 0-20% B in 20 min followed by 20-70% B in 100 min at a flow rate of 50 µl/min. UV detection was at 214 nm.

Desalting of the digestion mixtures before matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) MS analysis was done on a 0.8 mm × 2.0 mm Vydac C_4 pre-column (5 µm, 30 nm; LC-Packings) at a flow rate of 50 µl/min. Peptide mixtures were injected on to the column, pre-equilibrated with 0.1% TFA and eluted with 200 µl of 60% (v/v) ACN in 0.1% TFA.

SDS/PAGE and protein and peptide sequencing

SDS/PAGE was performed in 15% (w/v) gels by the method of Laemmli [26]. Before SDS/PAGE, RP fractions from interferon separation were dried in a vacuum centrifuge, dissolved in 10 µl of reducing sample buffer and boiled for 4 min. Sequence analysis was performed on an Applied Biosystems 494A Procise* sequencer (Perkin Elmer). For N-terminal sequencing, interferons from RP-HPLC fractions were further separated by SDS/PAGE [15% (w/v) gel] [26] followed by electroblotting on a ProBlott* membrane (Applied Biosystems, Perkin Elmer) in 10 mM Caps (pH 11)/10% (v/v) methanol with a constant potential of 50 V for 165 min [27]. The protein bands were cut out after being made visible with Coomassie Brilliant Blue and loaded on the sequencer.

MS

MALDI–TOF MS was performed on a Biflex* TOF instrument (Bruker-Franzen Analytik, Bremen, Germany) equipped with a nitrogen laser operating at 337 nm. RP-HPLC-isolated peptides or desalted digestion mixtures were analysed in the positive-ion

reflector mode by using a saturated solution of α-CCA in acetone as the matrix [28]. Complete digestion mixtures without desalting were diluted 1:5 with 0.1% TFA before depositing 1 µl of the sample on the target. The matrix used for these analyses was a saturated solution of α-CCA in acetone and 10 mg/ml nitrocellulose in acetone, mixed in a ratio of 3:1 [29]. To remove the interfering buffer salts before MALDI–TOF-analysis, the dried digestion mixture on the sample probe was washed with 0.1% TFA. The mass spectra were externally calibrated with angiotensin II and bovine insulin as standard peptides (Sigma).

Electrospray mass spectra were collected with an API 300 triple-quadrupole mass spectrometer (Perkin-Elmer Sciex Instruments, Thornhill, Ontario, Canada). RP fractions of interferons were either analysed directly, or dried and then dissolved in 50% (v/v) methanol in water containing 0.5% (v/v) acetic acid to a final concentration of 2 pmol/µl, and directly injected into the ESI-mass spectrometer with a syringe pump at a flow rate of 5 µl/min. The mass spectrometer was calibrated with a polypropylene glycol mixture as supplied by the instrument manufacturer.

Database searches

Database searches after MALDI analysis were made with either MASS SEARCH, MS-FIT or MOWSE.

Capillary isoelectric focusing

Capillary isoelectric focusing was performed on a BioFocus*3000 Capillary Electrophoresis System (Bio-Rad) with a 75 µm × 24 cm coated capillary (BioFocus; Bio-Rad) and internal calibration with BioMark synthetic pI markers (pI values 5.3, 6.4, 7.4, 8.4 and 10.4; Bio-Rad). The capillary was thermostatically controlled to 15 °C and the samples and reagents to 10 °C. RP fractions from interferon purification were dried in a vacuum centrifuge and dissolved in either CE-IEF Bio-Lyte*3-10 Ampholyte solution or 2% Ampholine*5-8 solution. Focusing was done for 4 min at +15 kV with 20 mM H_3PO_4 in the inlet buffer reservoir and 40 mM NaOH in the outlet buffer reservoir. Cathodic mobilization was performed with CE-IEF Mobilizer in the outlet buffer reservoir for 40 min at +15 kV. Detection was at 280 nm.

Proteolytic digestions

In-solution digestions

For in-solution digestions RP fractions were dried in a vacuum centrifuge and dissolved in 50 µl of 0.1 M Tris/HCl (pH 9.2)/10% (v/v) ACN. Endoproteinase LysC (0.1 µg; Wako Chemicals GmbH, Neuss, Germany) was added and the digestion mixture was incubated overnight at 37 °C. Disulphide bridges were reduced before desalting or RP separation by adding 2.5 µl of 28 mM DTT solution to the digestion mixture and incubating for 10 min at 60 °C. Digestions with trypsin (0.1 µg, Sequencing Grade Modified Trypsin, pig or bovine; Promega, Madison, WI, U.S.A.) or endoproteinase GluC (0.1 µg. Sequencing Grade Endoproteinase GluC; Promega) were done with the same method as above except that the digestion buffer was 0.1 M NH_4HCO_3/10% (v/v) ACN.

In-gel digestions

In-gel digestions were done by the method of Rosenfeld et al. [30], with some modifications. After staining with Coomassie Brilliant Blue and destaining with 30% (v/v) methanol, the appropriate interferon bands were excised from the gel and cut

into small pieces. The gel pieces were washed twice with 0.5 M Tris/HCl (pH 9.2)/50% (v/v) ACN for 45 min at 37 °C to remove SDS and Coomassie stain from the gel. The gel pieces were then dried for 30 min in a vacuum centrifuge. Gel pieces were re-wetted and the proteins were reduced by the addition of 50 μl of 14 mM DTT in 0.1 M Tris/HCl, pH 9.2, and incubation at 60 °C for 45 min. The excess DTT solution was removed and the gel pieces were again dried for 30 min in the vacuum centrifuge. The gel pieces were re-wetted with 10 μl of 0.1 M Tris/HCl (pH 9.2)/10% (v/v) ACN/0.04 μg/μl endo-proteinase LysC. After 10 min, 40 μl of 0.1 M Tris/HCl (pH 9.2)/10% (v/v) ACN was added. The digestion mixture was incubated overnight at 37 °C and the peptides generated were extracted twice with 150 μl of 0.1% TFA/60% (v/v) ACN for 30 min at 37 °C. The peptide extracts were combined and the excess ACN was removed in a vacuum centrifuge before RP-HPLC.

On-membrane digestions

On-membrane digestions were made by the method of Gharah-daghi et al. [31], with some modifications. The appropriate stained interferon bands were excised from the membrane and washed first with 0.1% (v/v) triethylamine in methanol to remove the Coomassie stain, followed by two washes with methanol. The washed bands were dried for 15 min in the vacuum centrifuge; the dried membrane pieces were then covered with digestion buffer and incubated overnight at 37 °C. The digestion buffer for endoproteinase LysC was 0.1 M Tris/HCl (pH 9.2)/10% (v/v) ACN/1% (v/v) octyl glucopyran-oside/4 ng/μl endoproteinase LysC; that for trypsin and endo-proteinase GluC was 50 mM NH_4HCO_3/10% (v/v) ACN/1% (v/v) octyl glucopyranoside/10 ng/μl enzyme. Disulphide bridges were reduced by the addition of 2.5 μl of 28 mM DTT to the digestion mixture and incubation at 60 °C for 10 min. The supernatant was withdrawn and analysed directly by MALDI-TOF. To extract any remaining peptides on the membrane, a wash with 50 μl of 60% (v/v) ACN for 30 min at 37 °C was performed. This extraction solution was also analysed by MALDI-TOF.

Glycopeptide analysis

The likely glycopeptides from IFN-α14 were subjected to MALDI-TOF and sequence analysis before, and to MALDI-TOF analysis after, the removal of N-glycans. Removal of the N-glycans was performed with recombinant N-glycosidase F (Boehringer Mannheim, Germany) by overnight hydrolysis at 37 °C in 50 mM NH_4HCO_3 with 0.2 unit of the enzyme.

RESULTS

For identification and structural characterization, the components of the highly purified leucocyte IFN-α preparation were first fractionated by RP-HPLC and then (either directly or after further separation by SDS/PAGE) subjected to N-terminal sequencing, mass analysis, pI determination and peptide mass mapping.

A typical RP-HPLC chromatogram obtained from the highly purified interferon preparation is shown in Figure 1(A) and the SDS/PAGE results from its collected fractions in Figure 1(B). (The shoulder peak for fraction 8 is termed 8b in Tables 1, 2 and 5.) Because most of the interferon components could not be completely separated by RP-HPLC, they were further purified by SDS/PAGE before N-terminal sequencing as well as in-gel and on-membrane digestions. For fractions with multiple bands in
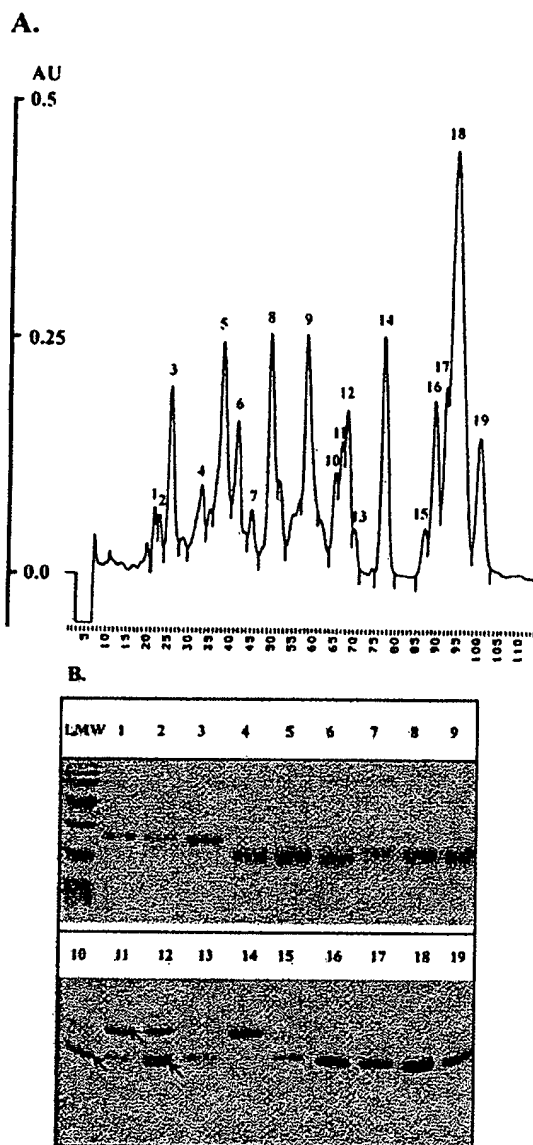

Figure 1   RP-HPLC separation of the highly purified IFN-α preparation and results of SDS/PAGE of collected fractions

(A) Interferons were separated by using a 2.0 mm × 150 mm Delta-Pak® HPI C$_4$ column (5 μm, 30 nm) and a linear gradient of 40–55% solvent B in 150 min. Solvent A was 0.1% TFA and solvent B was 0.075% TFA in ACN. The flow rate was 150 μl/min and UV detection was at 214 nm. Fractions were collected as indicated by the bars in the chromatogram. (B) Proteins were separated in a 15% (w/v) polyacrylamide gel and revealed by staining with Coomassie Brilliant Blue. The lane labelled LMW contained low-molecular-mass markers (Pharmacia), with molecular masses of 14.4, 20.1, 30.0, 43.0, 67.0 and 94.0 kDa.

SDS/PAGE, the band used for N-terminal sequencing is marked with an arrow in Figure 1(B). The results from N-terminal sequencing of the various components are shown in Table 1. Comparison of these sequences with the existing protein sequence data either exactly defined the IFN-α subtype or limited it to a few possibilities.

Table 1 The N-terminal sequences (one-letter codes) obtained for components in the human leucocyte IFN-α drug

The RP fractions are those shown in Figure 1(A). Abbreviation: C*, cysteine as an acrylamide derivative.

RP fraction	N-terminal sequence (positions 1–35; C = C*)	Possible IFN-α type
1	C N L S Q T H S L N N R R T L M L M A Q M R R – S P F S C L · D R · D	14a, 14b or 14c
2	C N L S Q T H S L N N R R T L M L M A Q M R R – S L F S C L · D R · D	14a, 14b or 14c
3	C N L S Q T H S L N N R R T L M L M A Q M R R – S L F S C L · D R · D	14a, 14b or 14c
4	C D L P Q T H S L G S R R T L · L L A Q M R R – S L F S C L · D R · D	2b or 2c
5	C D L P Q T H S L G S R R T L · L L A Q M R R – S L F S C L · D R · D	2b or 2c
6	C D L P Q T H S L G S R R A L · L L A Q M G R – S L F S C L · D R · D	2b or 2c
7	C D L P Q T H S L G N R R A L · L L A Q M G R – S L F S X L · D R X D	17a, 17b, 17c, 17d, 21a or 21b
8	C D L P Q T H S L G N R R A L · L L A Q M G R – S L F S C L · D R · D	17a, 17b, 17c, 17d, 21a or 21b
9	C D L P Q T H S L G N R R A L · L L A Q M G R – S L F S X L · D R · D	4a, 4b or 16
10	C D L P Q T H S L G N R R A L · L L A Q M G R – S L F S X L · D R · D	10a
11	C D L P Q T H S L G N R R A L · L L A Q M G R – S L F S X L · D R · D	17a, 17b, 17c, 17d, 21a or 21b
12	C D L P Q T H S L G R R R A L · L L A Q M G R – S L F S C L · D R · D	8a, 8b or 8c
13	C D L P Q T H S L G N R R A L · L L A Q M R R – S L F S C L · D R · D	17b, 17c, 17d, 21a or 21b
14	C D L P Q T H S L G D R R T L · L X A Q M R R – S L F S C L · D R · D	7a, 7b or 7c
15	C D L P Q T H S L G D R R T L M L L A Q M R S – S L F S C L · D R · D	8a, 8b or 8c
16	C D L P E T H S L D N R R T L M L L A Q M R S – X P X X C M K D R H D	1a or 1b
17	C D L P E T H S L D N R R T L M L L A Q M R S – S S S S C L · D R · D	1a or 1b
18	C D L P E T H S L D N R R T L M L L A Q M R S – S S S S C M · D R · D	1a or 1b
19	C D L P E T H S L D N R R T L M L L A Q M R S – S S S S C M · D R · D	1a or 1b

To identify the components further, they were subjected to peptide mass mapping. Proteolytic digestions were mainly made in solution with endoproteinase LysC, followed by reduction to separate peptides interconnected by disulphide bridges. These peptides could not otherwise be recognized by the existing mass search programs. The conventionally used methods for reducing, alkylating and desalting proteins before proteolytic digestion were also tried but the recoveries of the alkylated interferons were low. For those RP fractions giving multiple bands in SDS/PAGE, the peptide mass mapping was also performed from individual bands by in-gel or on-membrane digestion. The disadvantages of these methods are that especially large and/or hydrophobic peptides are poorly recovered and the peptide extract obtained from in-gel digestion is not directly compatible for MALDI-TOF analysis.

Mass searches were made with three different programs, each of which has its benefits for certain applications. Program MASS SEARCH gave the best results with peptides from an interferon mixture because it indicates the significance of the relative scores. By using this program and the results from N-terminal sequences, we ascertained that RP fractions 10, 11, 12 and 13 (Figure 1A) contained IFN-α17, IFN-α8, IFN-α17 and IFN-α7 respectively. The MOWSE mass search program was the most useful if the digestion mixture also contained fragments resulting from non-specific cleavages. When fragments contained oxidized residues, the program MS-FIT was the most useful. One problem with mass searches is that existing databases contain the proteins in their precursor form. Thus the N-terminal peptides of the mature proteins usually remain unrecognized. For this reason the relatively large and characteristic N-terminal peptides of the IFN-α subtypes obtained with endoproteinase LysC were useless in mass mapping. We used a combination of all these programs to identify the IFN-α subtypes in RP fractions 1–19 (Figure 1A); the results are shown in Table 2. All data presented in Table 2 were obtained from in-solution digestions followed by reduction and peptide separation by RP-HPLC. Figure 2 shows as an example the peptides obtained from RP fraction 3 (Figure 1A) and the MALDI-TOF results from these peptides; details of the peaks are given in Table 3. Because most of the interferons could not be baseline-separated by RP-HPLC, the mass maps also contain peptides from the adjacent components. Some non-specific cleavages were also observed, mainly Lys-Pro- and some Arg-Xaa.

Finally, the allelic or sequence variants of each subtype were identified by using data obtained from N-terminal sequencing, peptide mass mapping and mass analysis. As can be seen from Table 4, each variant contains characteristic amino acids [19], which leads to characteristic mass maps for each variant. IFN-α10a was the only variant that could be identified directly from the N-terminal sequence. Peptide mass mapping then identified the rest of the variants except for RP fraction 13 (Figure 1A), which was identified by electrospray ionization MS (ESI-MS) to be IFN-α7a.

All other RP fractions (Figure 1A) were also subjected to ESI-MS analysis without prior proteolytic digestion; the results are shown in Table 5. The masses correspond to those predicted within the error limits of our ESI-MS method (±2 Da for interferon-sized proteins) except for those known or suggested to be oxidized or glycosylated. Most of the interferons were eluted in RP-HPLC as two or more peaks. The ESI-MS results show that the peaks eluted earlier for the same subtype or variant differ in molecular mass by $+16$ or $+32$ Da from the component eluted later with the expected mass, suggesting that the earlier peaks represent oxidized forms of the protein. With IFN-α1a, the last two peaks had the expected mass, suggesting that they had

Table 2 Results from MALDI–TOF mass maps for RP fractions of IFN-α subtypes

The RP fractions are those shown in Figure 1(A). All data were obtained from in-solution digestions followed by reduction. The molecular masses shown with a superscript 'ox' differ by +16 or +32 Da from the calculated one, suggesting possible methionine oxidation. The molecular masses shown in bold correspond to the expected peptides of the adjacent subtypes; those shown in italic are unrecognized peptides (non-specific cleavages).

RP fraction	Obtained peptide masses used for database search (endoproteinase LysC digest) (Da)	Best fit for
1	1605.0, 2413.5, 2425.4, 2917.1, 3045.4, 4277.3ox	IFN-α14 precursor
2	2413.5, 2425.5, 2917.3, 3045.3, 3423.7, 4263.1	IFN-α14 precursor
3	727.9, 1604.9, 2413.6, 2424.7, 2916.3, 3044.6, 3243.8, 4260.9	IFN-α14 precursor
4	1030.7, 1113.4, 1344.7, 1808.9, 2226.6, 2413.8, 2425.1, 2460.3, 2917.3, 3045.7, 3302.8, 3489.9, 3617.9, 3706.3, 3747.2	IFN-α2 precursor
5	1030.9, 1344.5, 1450.7, 2226.2, 2460.2	IFN-α2 precursor
6	1029.8, 1246.1, 1344.4, 1450.1, 2226.1, 2357.1, 2459.5, 3489.3, 3576.6, 3617.8	IFN-α2 precursor
7	1030.1, 1059.5, 1245.4, 1284.0, 1344.3, 1374.7, 1476.1, 1574.9, 2226.4, 2341.5, 2406.8, 2460.5, 2588.8, 4108.7ox, 4382.1, 4455.8	IFN-α21 precursor
8	1245.3, 1344.8, 1574.4, 2225.5, 2340.6, 2350.1, 2406.0, 2459.7, 2587.9, 4091.3	IFN-α21 precursor
8b	981.0, 1031.3, 1297.5, 1345.4, 1451.6, 1575.7, 1603.8, 1954.8, 2109.9, 2227.4, 2308.5, 2342.4, 2366.4, 2370.2, 2407.9, 2461.2, 2566.6, 2918.6, 3045.8, 3305.3, 3449.4, 3481.9, 3708.3, 3723.2, 3852.3, 3884.9, 4148.9, 4220.3, 4327.8ox, 5884.9	IFN-α4 precursor
9	1603.4, 1743.6, 2227.1, 2407.1, 2917.2, 3426.6, 3854.4, 4203.1, 4326.1, 5842.7	IFN-α10 precursor
10	1575.5, 1603.2, 2307.9, 2479.1, 3883.8	IFN-α17 precursor
11	1574.9, 1603.2, 1953.5, 2307.3, 2406.9, 2918.8, 3884.2	IFN-α8 precursor
11, 12 and 13	1574.8, 1603.0, 1954.2, 2307.0, 2407.2, 2478.4, 2822.5, 2916.2, 3045.4, 3441.0, 3884.2, 4357.0*	IFN-α17, 8 and 7 precursors
14	1345.0, 1575.3, 1953.8, 2406.8, 2918.0, 5779.4	IFN-α8 precursor
15	1603.0, 1949.3, 2148.6, 2384.7, 3594.4, 3722.7, 3851.8, 4236.5ox, 5676.4, 5843.8	IFN-α1 precursor
16	1464.9, 1602.2, 2383.9, 5838.2	IFN-α1 precursor
17	1464.2, 1602.5, 2385.0, 4203.8, 5846.0	IFN-α1 precursor
18	1602.8, 2383.8, 4200.8, 5842.9	IFN-α1 precursor
19	1465.0, 1603.6, 2384.7, 3594.0, 4201.6, 5844.9	IFN-α1 precursor

* All masses correspond to one of the identified subtypes (IFN-α 17, 8 or 7).
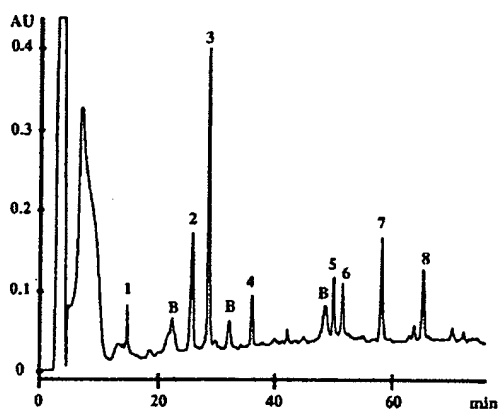
Figure 2 Peptide separation from in-solution digestion of RP fraction 3

Peptides obtained from the digestion of RP fraction 3 (Figure 1A) were separated on a Vydac 1.0 mm × 150 mm C₈ column (5 μm, 30 nm). Elution was performed with linear gradients of 0–20% B over 20 min, followed by 20–70% B over 100 min, at a flow rate of 50 μl/min. Solvent A was 0.1% TFA and solvent B was 0.075% TFA in ACN. UV detection was at 214 nm. The numbered peaks were collected and the molecular masses of these peptides were measured by MALDI–TOF-MS; these are listed in Table 3. Peaks labelled B were also found in the blank chromatogram.

Table 3 The observed and predicted molecular masses for peptides from in-solution digestion of RP fraction 3

The positions of the identified peptides from IFN-α14c are also shown. The presence of a glycan on the peptide in peak 2 was confirmed by sequence analysis (see the text). Abbreviation: +ox, the observed peptide also contained the oxidized form.

Peak	Molecular mass (Da)		Residues from IFN-α14c
	Observed	Predicted	
1	727.9	727.9	Arg161-Lys165
2	3243.9, 3257.4	1463.6 + glycan	Cys1-Lys31
3	2413.6	2413.5	Asp32-Lys50
4	1604.9,1620.9	1640.9 (+ox)	Tyr123-Lys134
5	3044.6, 3060.9	3044.5 (+ox)	Lys135-Lys160
6	2916.3, 2932.6	2916.4 (+ox)	Tyr136-Lys160
7	2424.7, 2440.8	2424.8 (+ox)	Ala51-Lys71
8	4260.9	4259.9	Phe85-Lys121

different conformations. Recombinant human IFN-α2b has previously been shown to be eluted in RP-HPLC as two peaks, where the component eluted earlier was a monomethionine sulphoxide form of the protein [32]. These oxidation and chromatographic conditions were also shown to change the conformation of recombinant human IFN-α2b as measured by CD spectrometry [32].

IFN-α2 is the only subtype shown to be O-glycosylated [9]; the glycosylation was also clearly demonstrated in our analysis. IFN-α14 is the only IFN-α subtype that contains potential N-glycosylation sites, one at Asn-2 and another at Asn-72. The results from the N-terminal sequence analysis (Table 1) clearly show that the first potential glycosylation site (Asn-2) in IFN-α14 does not contain a glycan, because the phenylthiohydantoin-Asn from position 2 was recovered from the protein in a molar amount corresponding to that of the analysed protein. To further characterize the glycosylation at the remaining potential glycosylation site (Asn-72), we isolated the putative glycopeptide from both an endoproteinase LysC digest and an endoproteinase GluC digest of IFN-α14. The endoproteinase LysC peptide, which should start with Asn-72, was first subjected to

Table 4   Possible variants for the IFN-α subtypes found from the highly purified IFN-α preparation together with their characteristic amino acids and LysC-peptides

In the entry for IFN-α8c, Δ(162–166) means that these residues are deleted.

IFN-α type	Amino acid residues characteristic of the presumed variant	Molecular mass(es) of the characteristic LysC-peptide(s) (Da)
1a	A^{114}	4202.8
1b	V^{114}	4230.9
2a	K^{23}, H^{34}	2656.2, 910.1 and 2226.3
2b	R^{23}, H^{34}	3576.3 and 2226.3
2c	R^{23}, R^{34}	3576.3 and 2245.4
4a	A^{51}, E^{114}	3853.3 and 4325.9
4b	T^{51}, V^{114}	3883.3 and 4295.9
7a	M^{132}, K^{159}, G^{161}	1604.9, 2822.2 and 628.8
7b	M^{132}, D^{159}, R^{161}	1604.9, 2950.4 and 727.9
7c	T^{132}, K^{159}, G^{161}	1574.8, 2822.2 and 628.8
8a	V^{98}, L^{99}, C^{100}, D^{101}, R^{161}	5789.4 and 415.5
8b	S^{98}, C^{99}, V^{100}, M^{101}, R^{161}	5779.4 and 415.5
8c	S^{98}, C^{99}, V^{100}, M^{101}, D^{161}, Δ(162–166)*	5779.4 and 133.1
10a	S^8, L^{83}	3426.1 and 4325.9
10b	T^8, P^{83}	3440.1 and 4325.9
14a	F^{152}, D^{159}, R^{161}	2950.4 and 727.9
14b	F^{152}, K^{159}, G^{161}	2822.2 and 628.8
14c	L^{152}, D^{159}, R^{161}	2916.4 and 727.9
17a	P^{34}, S^{55}, I^{161}	2267.4, 3883.3 and 684.9
17b	H^{34}, S^{55}, I^{161}	2307.4, 3883.3 and 684.9
17c	H^{34}, S^{55}, R^{161}	2307.4, 3883.3 and 727.9
17d	H^{34}, P^{55}, R^{161}	2307.4, 3893.3 and 727.9
21a	M^{96}	4108.7
21b	L^{96}	4090.6

Table 5   Identified protein components from the highly purified IFN-α drug together with the measured and calculated molecular masses and pI values for these proteins

The RP fractions are those shown in Figure 1(A). Abbreviations: cIEF, capillary isoelectric focusing: ^{gly}, known or suggested to be glycosylated; ^{ox}, molecular masses differed by +16 or +32 Da from the calculated one, suggesting possible methionine oxidation. Dashes indicate values not measured.

RP fraction	Identified IFN-α	Molecular mass (Da)		pI	
		Measured by ESI-MS	Calculated	Measured by cIEF	Calculated
1	14c	21780^{gly}	19704	–	6.76
2	14c	21780^{gly}	19704	–	6.76
3	14c	21765^{gly}	19704	6.12	6.76
4	2b	20285^{gly}	19265	5.99	6.28
5	2b	20211^{gly}	19265	5.58	6.28
6	2b	19921^{gly}	19265	6.01	6.28
7	21b	19322^{ox}	19308	–	6.28
8	21b	19307	19308	6.05	6.28
8b	4b	19374	19375	–	5.76
9	10a	19403	19402	6.18	5.80
10	17b	19309^{ox}	19295	5.90	5.40
11	8b	19495^{ox}	19480	5.65	5.01
12	17b	19297	19295	5.90	5.40
13	7a	19601	19602	–	6.14
14	8b	19481	19480	5.65	5.01
15	1a	19414^{ox}	19382	5.62	5.00
16	1a	19398^{ox}	19382	5.62	5.00
17	1a	19398^{ox}	19382	5.55	5.00
18	1a	19381	19382	5.55	5.00
19	1a	19382	19382	5.55	5.00

MALDI–TOF-analysis, yielding a molecular mass of 3552 Da. Analysis of the same peptide after one cycle of Edman degradation gave a mass of 1349.5 Da, which clearly shows that the peptide is glycosylated at its first residue (Asn-72). Further confirmation of the glycosylation site with N-glycosidase F was

not possible from the endoproteinase LysC peptide owing to the terminal position of Asn-72. Therefore we isolated the corresponding endoproteinase GluC peptide (from Met-60 to Glu-79) and confirmed its identity by N-terminal sequencing. The molecular mass of the peptide was measured both before (approx.

4128 Da) and after (2337.9 Da) the N-glycosidase F treatment, which showed that the peptide contains a glycan with a molecular mass of approx. 1800 Da.

For further characterization of different IFN-α subtypes we subjected most of them to capillary isoelectric focusing analysis; the results are shown in Table 5. As can be seen, the experimentally determined pI values differ from the theoretical ones calculated from the amino acid compositions. The reason for this is that the three-dimensional structure of a protein has an effect on the amino acid side chain pK_a values in the polypeptide chain [33]. Glycosylation of IFN-α2 and IFN-α14 also causes microheterogeneity in the capillary isoelectric focusing electropherograms of these proteins.

DISCUSSION

In the present paper we have demonstrated that at least nine different IFN-α subtypes are produced in normal human peripheral leucocytes in response to Sendai virus infection. The major subtypes are IFN-α1a and IFN-α2b, which comprise respectively approx. 30% and 20% of all IFN-α protein. The other main subtypes are IFN-α8b, IFN-α10a, IFN-α14c, IFN-α17b and IFN-α21b, whereas IFN-α4b and IFN-α7a represent minor components. Only one sequence variant was identified for each subtype, suggesting that the other gene sequences described represent rare allelic variants or mutations.

The starting material for this work was obtained by purification from the supernatant of a Sendai virus-induced leucocyte culture by two monoclonal antibodies that together bind approx. 90% of the total IFN activity produced by normal leucocytes (H. Tölö, unpublished work). The unbound IFN activity is composed of IFN-ω, IFN-β and the small amounts of IFN-γ known to be present in leucocyte culture supernatants [34,35]. However, losses of some minor IFN-α species during the purification cannot be excluded.

Seven IFN-α mRNA species have been identified in human peripheral blood leucocytes induced by Sendai virus [7]. The major mRNA species corresponded to IFN-α1, IFN-α2 and IFN-α4, whereas the minor ones were IFN-α5, IFN-α7, IFN-α8 and IFN-α14. The present study found that IFN-α4 was a minor subtype at the protein level and that IFN-α8 and IFN-α14 were relatively abundant subtypes. This suggests that there is not a strict correlation between the mRNA level and extracellular protein accumulation. However, in the present study we did not perform quantitative IFN-α mRNA analysis.

Most of the IFN-α subtypes characterized in the present paper have been identified previously in IFN-α preparations purified from leucocyte culture supernatants. The preliminary reports on the presence of IFN-α1 and IFN-α2b [10], IFN-α8b [11], IFN-α10 [10], IFN-α10a [11], IFN-α14 [10], IFN-α17b [11] and IFN-α21 [8,10] in leucocyte IFN-α are consistent with the present study. Only IFN-α2 has been characterized in more detail previously [9] and its major allelic variant IFN-α2b has been also identified at the protein level [12]. Recently the gene variants of four IFN-α species were identified in the genome of a large human population by PCR [21,22]. The identified gene variants were IFN-α7a, IFN-α8b, IFN-α14c and IFNα-21b, which are the same as those identified at the protein level in the present study.

We found IFN-α1a to be one of the major IFN-α species produced by human leucocytes in response to Sendai virus infection. This sequence variant was initially cloned from IFN-α produced by human leucocytes [36], whereas the other described variant, IFN-α1b, was cloned from a myeloblastoid cell line [37].

This suggests that the IFN-α1b variant represents a mutation not present in the normal human leucocyte genome.

IFN-α2b has been earlier shown to be O-glycosylated [9]; this was clearly demonstrated also in our analysis. IFN-α14 is the only IFN-α subtype that contains potential N-glycosylation sites, one at position Asn-2 and another at Asn-72. Our results indicate that only the latter site carries a carbohydrate with a molecular mass of approx. 1800 Da. It is noteworthy that the molecular mass determination of the whole proteins by ESI-MS indicated no post-translational modifications in the IFN-α subtypes other than the glycosylation of IFN-α2b and IFN-α14c. Detailed characterization of the carbohydrate structures of these glycosylated IFN-α species is currently under way.

REFERENCES

1 Cantell, K., Hirvonen, S., Kauppinen, H.-L. and Myllylä, G. (1981) Methods Enzymol. 78, 29–38

2 Finter, N. B. (1981/1982) Texas Rep. Biol. Med. 41, 175–178

3 Familletti, P. C., McCandliss, R. and Pestka, S. (1981) Antimicrob. Agents Chemother. 20, 5–9

4 Diaz, M. O. (1995) Semin. Virol. 6, 143–149

5 Todokoro, K., Kioussis, D. and Weissmann, C. (1984) EMBO J. 3, 1809–1812

6 Kontsek, P. (1994) Acta Virol. 38, 345–360

7 Hiscott, J., Cantell, K. and Weissmann, C. (1984) Nucleic Acids Res. 12, 3727–3746

8 Adolf, G. R., Kalsner, I., Maurer-Fogy, I., Ahorn, H. and Cantell, K. (1990) J. Interferon Res. 10 (Suppl. 1), S57

9 Adolf, G. R., Kalsner, I., Ahorn, H., Maurer-Fogy, I. and Cantell, K. (1991) Biochem. J. 276, 511–518

10 Pohl, G., Hellman, U., Lundgren, W. and Lundberg, S. (1994) J. Interferon Res. 14 (Suppl. 1). S63

11 DiPaola, M., Smith, T., Ogin, E., Marcotte, J.. Ferencz-Biro, K., Liao, M.-J. and Testa, D. (1994) J. Interferon Res. 14 (Suppl. 1), S72

12 DiPaola, M., Smith, T., Ferencz-Biro, K., Liao, M.-J. and Testa, D. (1994) J. Interferon Res. 14, 325–332

13 Evinger, M., Rubenstein, M. and Pestka, S. (1980) Ann. N. Y. Acad. Sci. 350, 17761–17764

14 Weck, P. K., Apperson, S., May, L. and Stebbing, N. (1981) J. Gen. Virol. 57, 233–237

15 Streuli, M., Hall, A., Boll, W., Stewart, W. E., Nagata, S. and Weissmann, C. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2848–2852

16 Ortlando, J. R., Herberman, R. B., Harvey, C., Osheroff, P., Pan, Y.-C., Keller, B. and Pestka, S. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 4926–4929

17 Sperber, S. J., Gocke, D. J., Haberzettl, C., Kuk, R., Schwartz, B. and Pestka, S. (1992) J. Interferon Res. 12, 363–368

18 Foster, G. R., Rodrigues, O., Ghouze, F., Schulte-Frohlinde, E., Testa, D., Liao, M.-J., Stark, G. R., Leadbeater, L. and Thomas, H. C. (1996) J. Interferon Cytokine Res. 16, 1027–1033

19 Allen, G. and Diaz, M. O. (1996) J. Interferon Res. 16, 181–184

20 Lee, N., Ni, D., Brissette, R., Chou, M., Hussain, M., Gill, D. S., Liao, M.-J. and Testa, D. (1995) J. Interferon Cytokine Res. 15, 341–349

21 Hussain, M., Gill, D. S. and Liao, M.-J. (1996) J. Interferon Res. 16, 853–859

22 Hussain, M., Gill, D. S., Liao, M.-J. and Testa, D. (1996) J. Interferon Res. 16, 523–529

23 Golovleva, I., Kandefer-Szerszen, M., Beckman, L. and Lundgren, E. (1996) Am. J. Hum. Genet. 59, 570–578

24 McKenna, R. M. and Oberg, K. E. (1997) J. Interferon Cytokine Res. 17, 141–143

25 Antonelli, G. (1995) J. Biol. Regul. Homeost. Agents 9, 123–131

26 Laemmli, U.K. (1970) Nature (London) 227, 680–685

27 Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038

28 Vorm, O., Roepstorff, P. and Mann, M. (1994) Anal. Chem. 66, 3281–3287

29 Jensen, O. N., Podtelejnikov, A. and Mann, M. (1996) Rapid Commun. Mass Spectrom. 10, 1371–1378

30 Rosenfeld, J., Capdevielle, J., Guillemot, J. G. and Ferrara, P. (1992) Anal. Biochem. 203, 173–179

31 Gharahdaghi, F., Kirchner, M., Fernandez, J. and Mische, S. M. (1996)
    Anal. Biochem. **233**, 94–99

32 Gitlin, G., Tsarbopoulos, A., Patel, S. T., Sydor, W., Pramanik, B. N., Jacobs, S.,
    Westreich, L., Mittelman, S. and Bausch, J. N. (1996) Pharmaceut. Res. **13**,
    762–769

33 Liu, X., Sosic, Z. and Krull, I. S. (1996) J. Chromatogr. A**735**, 165–190

34 Adoll, G. A. (1990) Virology **175**, 410–417

35 Hua, J., Liao, M.-J. and Rashidbaigi, A. (1996) J. Leukocyte Biol. **60**, 125–128

36 Nagata, S., Mantei, N. and Weissmann, C. (1980) Nature (London) **287**, 401–408

37 Goeddel, D. V., Leung, D. W., Dull, T. J., Gross, M., Lawn, R. M., McCandliss, R.,
    Seeburg, P. H., Ullrich, A., Yelverton, E. and Gray, P. W. (1981) Nature (London)
    290, 20–26

PCR#1

Reaction #1

Primer #1 ⟶ 75
⟵ Primer #2

~~IFN-α 21a Gene~~
IFN-α 21b Gene

Reaction #2

75
Primer #4 ⟵
Primer #3 ⟶

IFN-α 2c Gene

↓ PCR Amplification

75

~~α 21a DNA Fragments~~
α 21b DNA Fragments

75

α 2c DNA Fragments

PCR#2

IFN-α21b
~~IFN-α21a~~   75   IFN-α2c
Primer #1 ⟶   ⟵ Primer #4   Reaction #1
and #2 Fragments

↓

IFN-α21b
~~IFN-α21a~~   75   IFN-α2c   HY1
(IFN-α Hybrid 1)
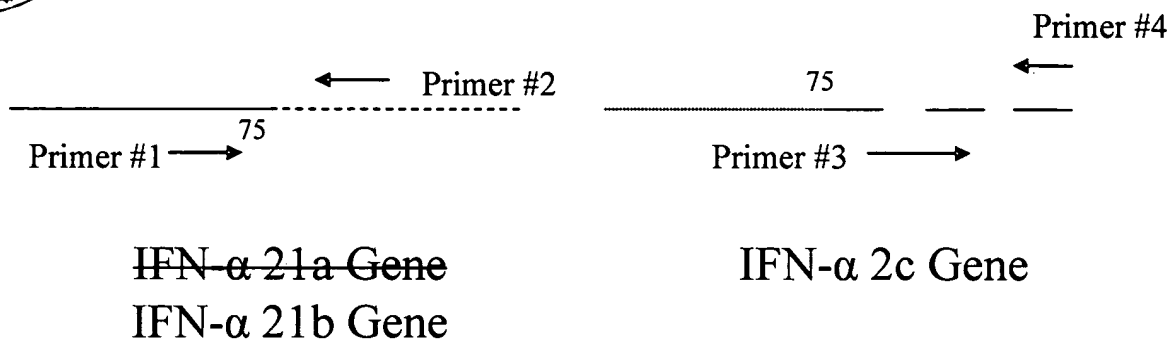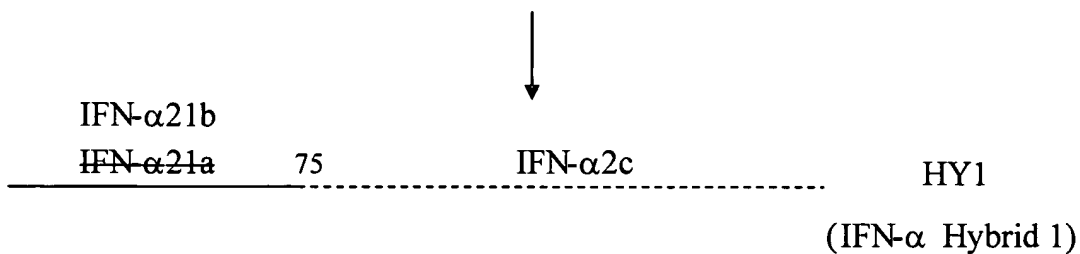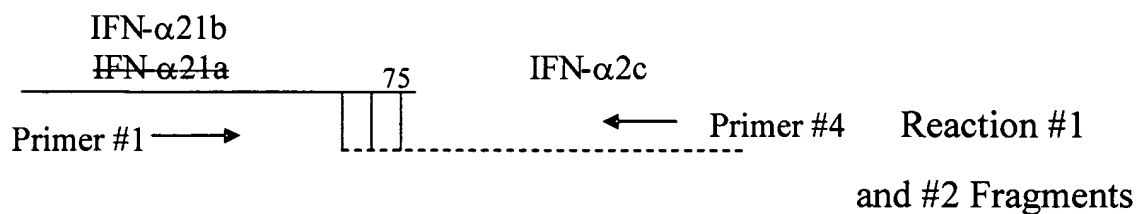
Construction of Hybrid #1

FIG. 1A

PCR#1

Reaction #1   Reaction #2
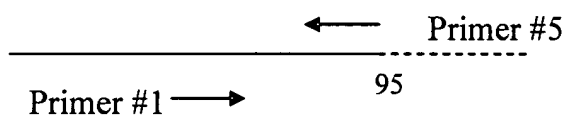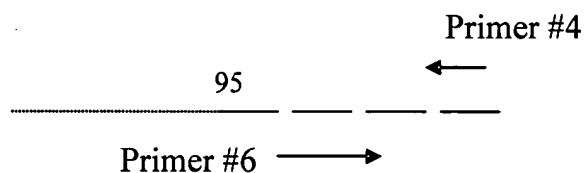
IFN-α 21a Gene
IFN-α 21b Gene   IFN-α 2c Gene

PCR Amplification

α 21a DNA Fragments
α 21b DNA Fragments   α 2c DNA Fragments

PCR#2

IFN-α21b
IFN-α21a
IFN-α2c

Primer #1 ⟶   Primer #4   Reaction #1

and #2 Fragments

IFN-α21b
IFN-α21a   IFN-α2c   HY2

(IFN-α Hybrid 2)
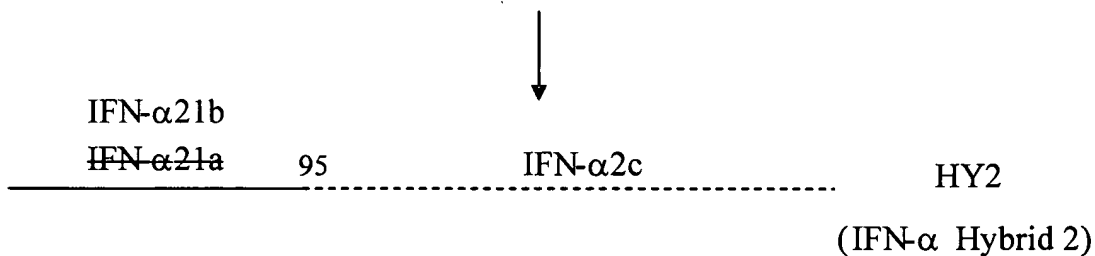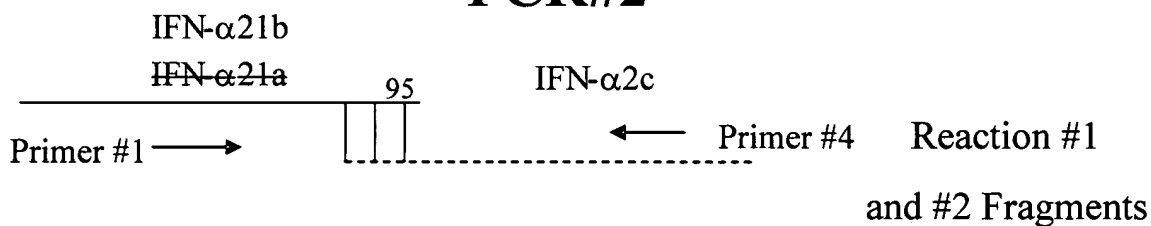
Construction of Hybrid #2

FIG. 1B

PCR#1

Reaction #1   Reaction #2

Primer #7

← Primer #5   95   ←

Primer #1 ⟶   95   Primer #6 ⟶

IFN-α 2c Gene   ~~IFN-α 21a Gene~~
                                 IFN-α 21b Gene

↓ PCR Amplification

____95____   ____95____

α 2c DNA Fragments   ~~α 21a DNA Fragments~~
                                 α 21b DNA Fragments

PCR#2

IFN-α21b
~~IFN-α21a~~

IFN-α2c   95
Primer #1 ⟶   ← Primer #7   Reaction #1
                                               and #2 Fragments

↓

IFN-α21b
IFN-α2c   95   ~~IFN-α21a~~   HY3
                                              (IFN-α Hybrid 3)
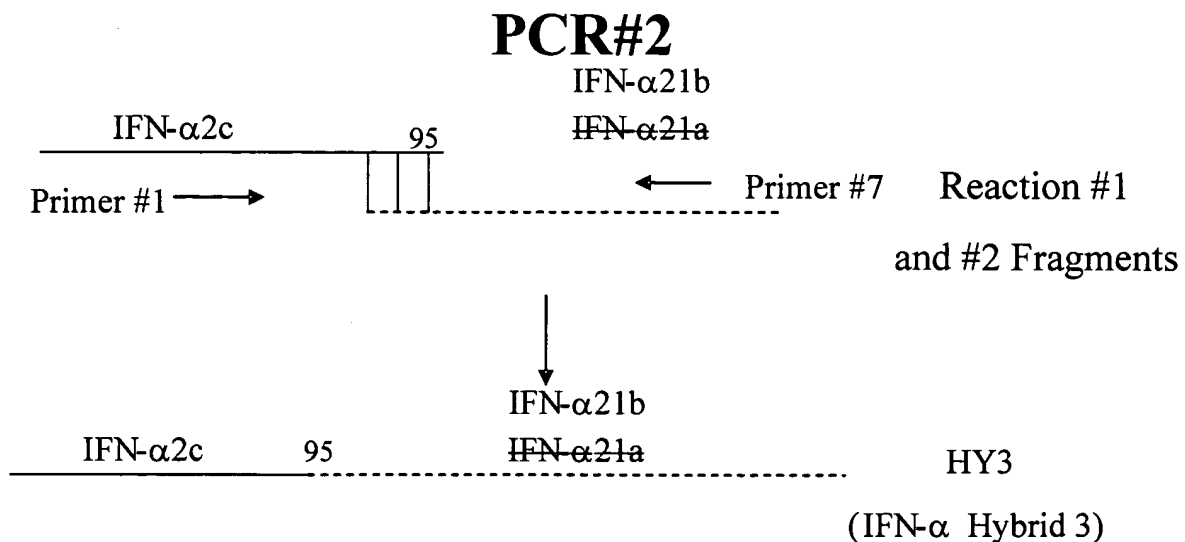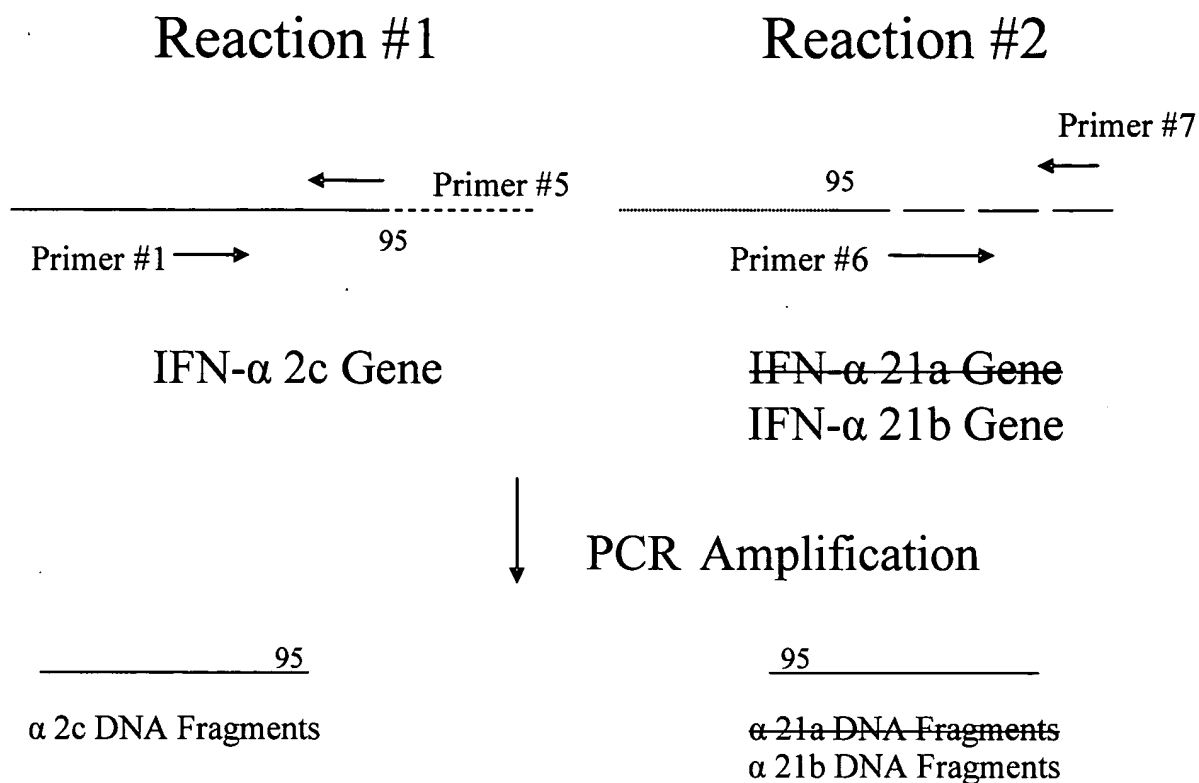
Construction of Hybrid #3

FIG. 1C
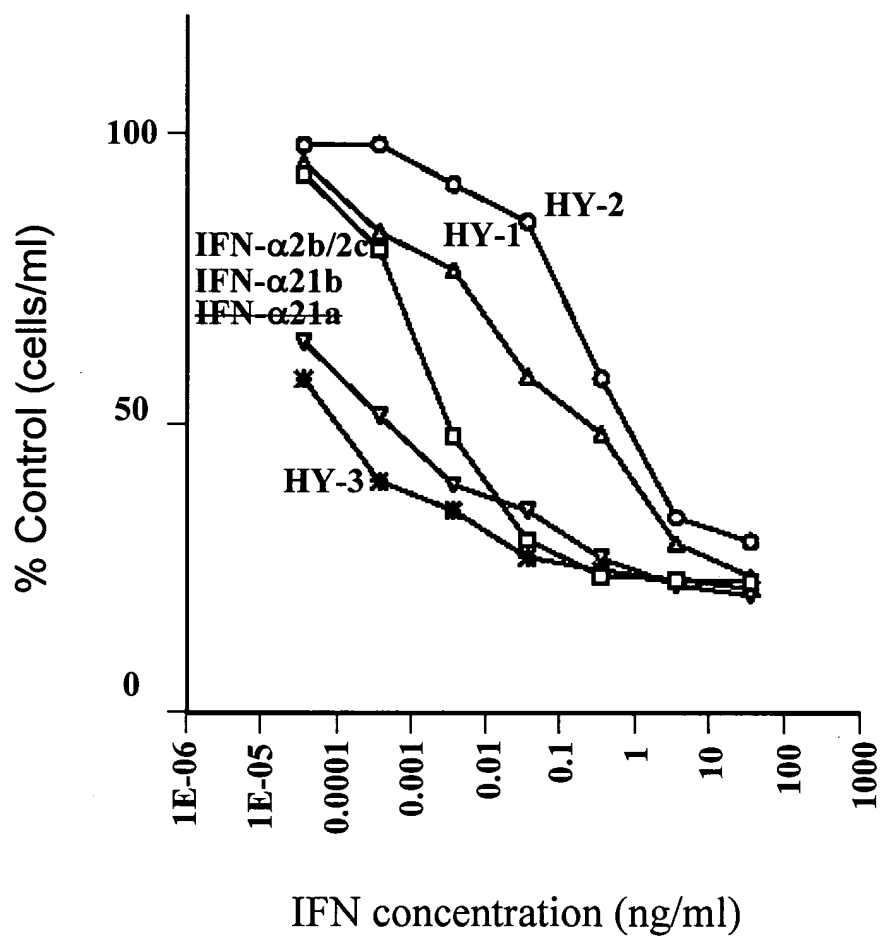
FIG. 2A
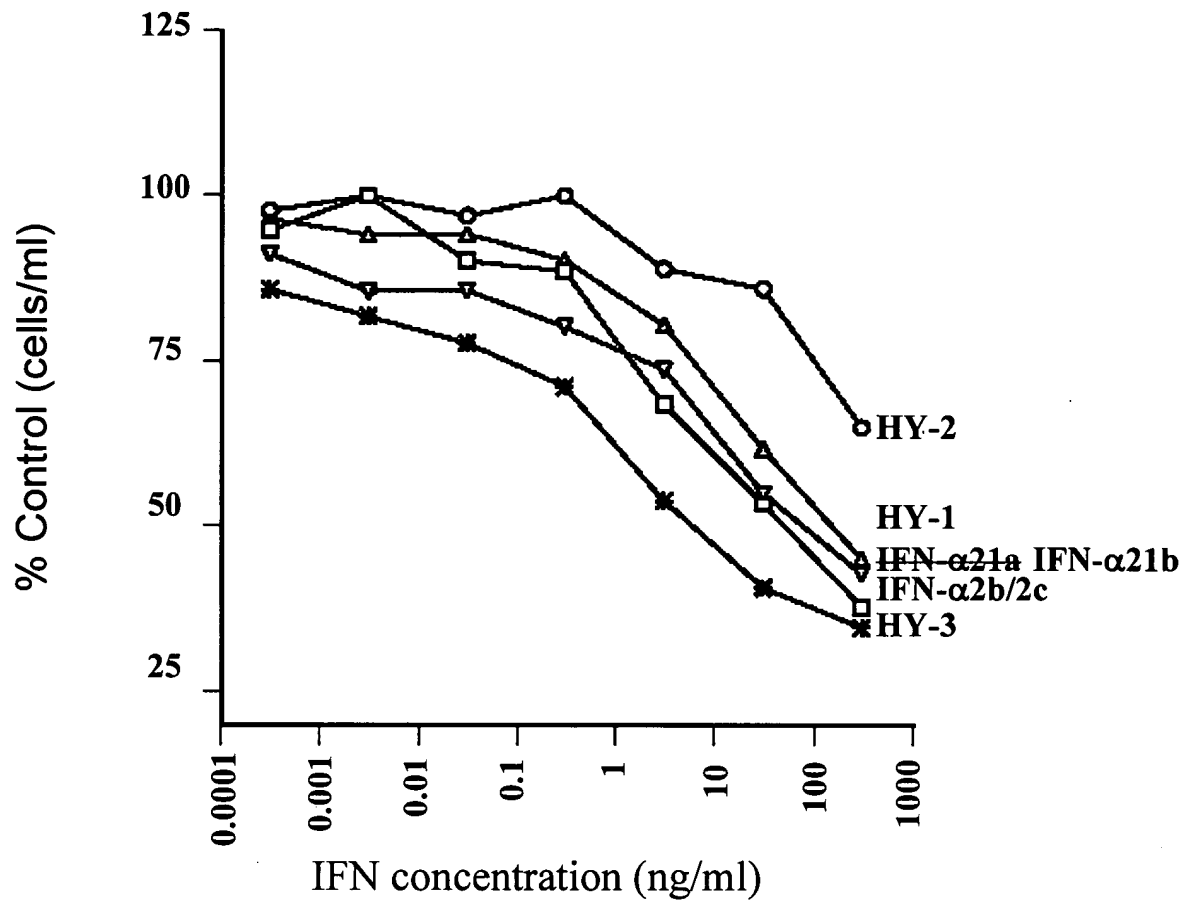
FIG. 2B
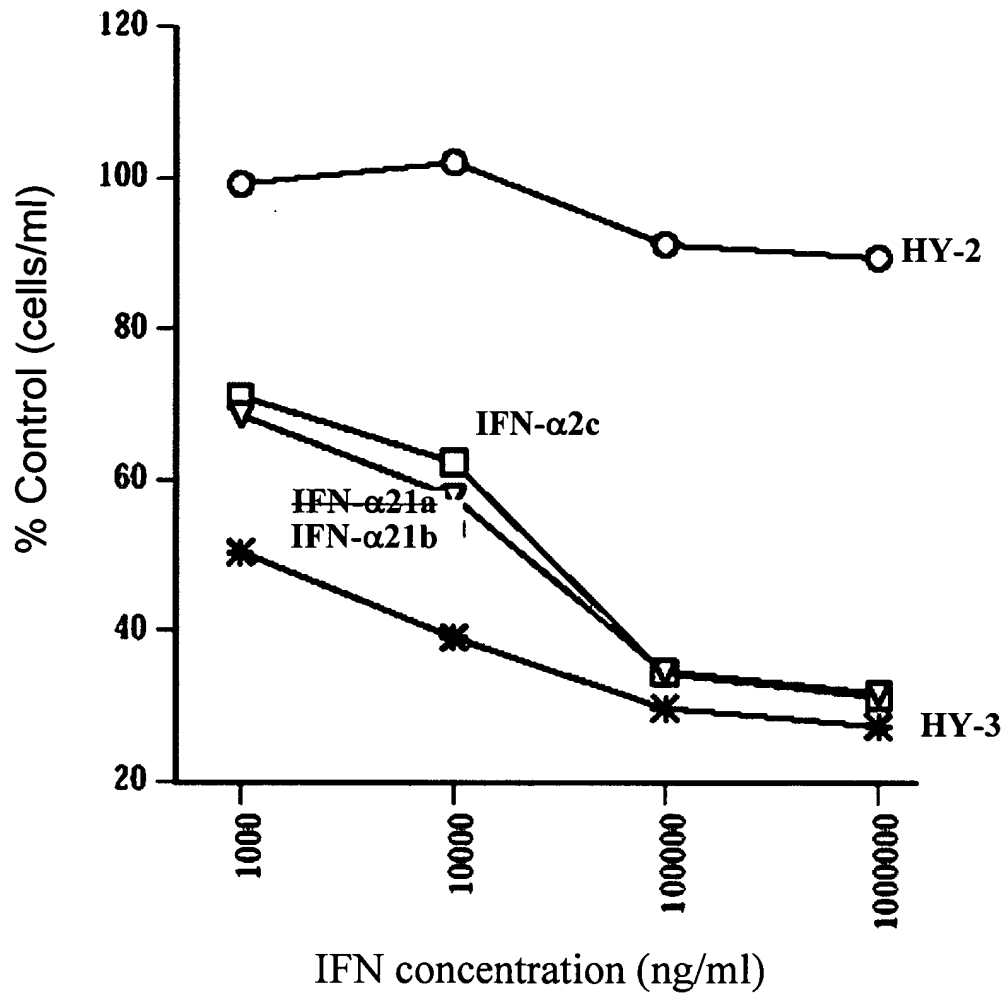
FIG. 2C
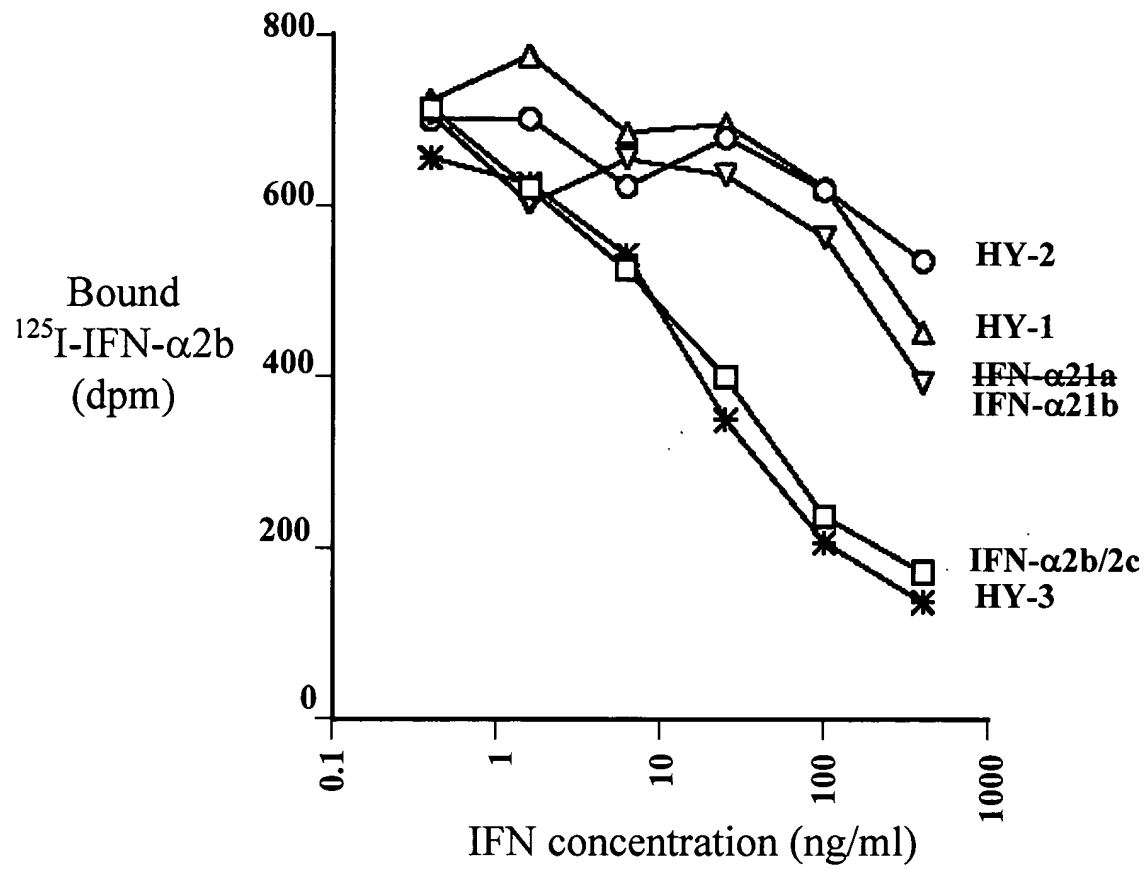
FIG. 4A
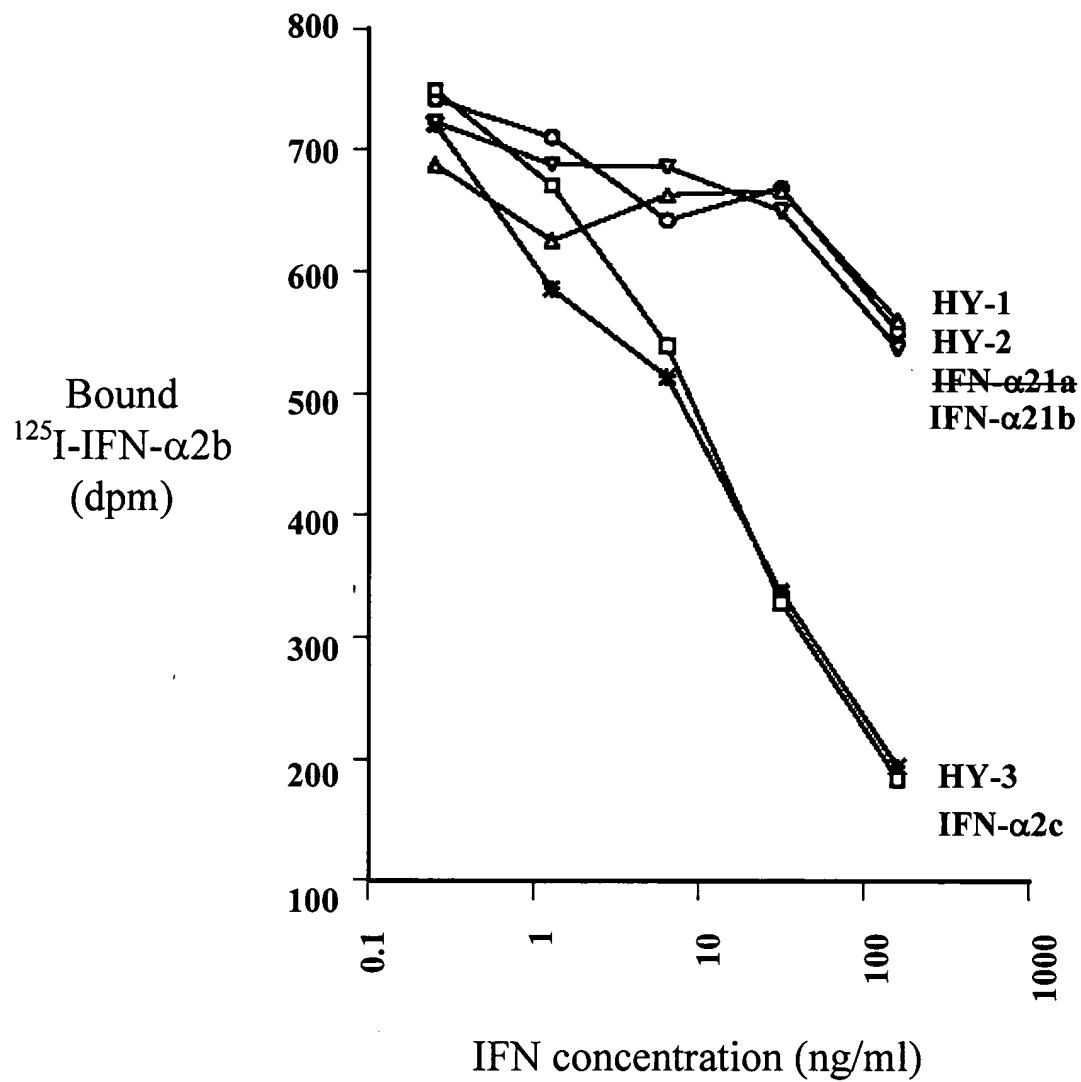
FIG. 4B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Alpha 2c	C	D	L	P	Q	T	H	S	L	G	S	R	R	T	L	M	L
Alpha 21ab	C	D	L	P	Q	T	H	S	L	G	N	R	R	A	L	I	L
HY-1	C	D	L	P	Q	T	H	S	L	G	N	R	R	A	L	I	L
HY-2	C	D	L	P	Q	T	H	S	L	G	N	R	R	A	L	I	L
HY-3	C	D	L	P	Q	T	H	S	L	G	S	R	R	T	L	M	L

	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Alpha 2c	L	A	Q	M	R	R	I	S	L	F	S	C	L	K	D	R	R
Alpha 21ab	L	A	Q	M	G	R	I	S	P	F	S	C	L	K	D	R	H
HY-1	L	A	Q	M	G	R	I	S	P	F	S	C	L	K	D	R	H
HY-2	L	A	Q	M	G	R	I	S	P	F	S	C	L	K	D	R	H
HY-3	L	A	Q	M	R	R	I	S	L	F	S	C	L	K	D	R	R

	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
Alpha 2c	D	F	G	F	P	Q	E	E	F	*	G	N	Q	F	Q	K	A
Alpha 21ab	D	F	G	F	P	Q	E	E	F	D	G	N	Q	F	Q	K	A
HY-1	D	F	G	F	P	Q	E	E	F	D	G	N	Q	F	Q	K	A
HY-2	D	F	G	F	P	Q	E	E	F	D	G	N	Q	F	Q	K	A
HY-3	D	F	G	F	P	Q	E	E	F	*	G	N	Q	F	Q	K	A

	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
Alpha 2c	E	T	I	P	V	L	H	E	M	I	Q	Q	I	F	N	L	F
Alpha 21ab	Q	A	I	S	V	L	H	E	M	I	Q	Q	T	F	N	L	F
HY-1	Q	A	I	S	V	L	H	E	M	I	Q	Q	T	F	N	L	F
HY-2	Q	A	I	S	V	L	H	E	M	I	Q	Q	T	F	N	L	F
HY-3	E	T	I	P	V	L	H	E	M	I	Q	Q	I	F	N	L	F

	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
Alpha 2c	S	T	K	D	S	S	A	A	W	D	E	T	L	L	D	K	F
Alpha 21ab	S	T	K	D	S	S	A	T	W	E	Q	S	L	L	E	K	F
HY-1	S	T	K	D	S	S	A	A	W	D	E	T	L	L	D	K	F
HY-2	S	T	K	D	S	S	A	T	W	E	Q	S	L	L	E	K	F
HY-3	S	T	K	D	S	S	A	A	W	D	E	T	L	L	D	K	F

	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102
Alpha 2c	Y	T	E	L	Y	Q	Q	L	N	D	L	E	A	C	V	I	Q
Alpha 21ab	S	T	E	L	N	Q	Q	L	N	D	L	E	A	C	V	I	Q
HY-1	Y	T	E	L	Y	Q	Q	L	N	D	L	E	A	C	V	I	Q
HY-2	S	T	E	L	N	Q	Q	L	N	D	L	E	A	C	V	I	Q
HY-3	Y	T	E	L	Y	Q	Q	L	N	D	L	E	A	C	V	I	Q

	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
Alpha 2c	G	V	G	V	T	E	T	P	L	M	K	E	D	S	I	L	A
Alpha 21ab	E	V	G	V	E	E	T	P	L	M	N	V	D	S	I	L	A
HY-1	G	V	G	V	T	E	T	P	L	M	K	E	D	S	I	L	A
HY-2	G	V	G	V	T	E	T	P	L	M	K	E	D	S	I	L	A
HY-3	E	V	G	V	E	E	T	P	L	M	N	V	D	S	I	L	A

	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136
Alpha 2c	V	R	K	Y	F	Q	R	I	T	L	Y	L	K	E	K	K	Y
Alpha 21ab	V	K	K	Y	F	Q	R	I	T	L	Y	L	T	E	K	K	Y
HY-1	V	R	K	Y	F	Q	R	I	T	L	Y	L	K	E	K	K	Y
HY-2	V	R	K	Y	F	Q	R	I	T	L	Y	L	K	E	K	K	Y
HY-3	V	K	K	Y	F	Q	R	I	T	L	Y	L	T	E	K	K	Y

	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153
Alpha 2c	S	P	C	A	W	E	V	V	R	A	E	I	M	R	S	F	S
Alpha 21ab	S	P	C	A	W	E	V	V	R	A	E	I	M	R	S	F	S
HY-1	S	P	C	A	W	E	V	V	R	A	E	I	M	R	S	F	S
HY-2	S	P	C	A	W	E	V	V	R	A	E	I	M	R	S	F	S
HY-3	S	P	C	A	W	E	V	V	R	A	E	I	M	R	S	F	S

	154	155	156	157	158	159	160	161	162	163	164	165	166	
Alpha 2c	L	S	T	N	L	Q	E	S	L	R	S	K	E	
Alpha 21ab	L	S	K	I	F	Q	E	R	L	R	R	K	E	
HY-1	L	S	T	N	L	Q	E	S	L	R	S	K	E	
HY-2	L	S	T	N	L	Q	E	S	L	R	S	K	E	
HY-3	L	S	K	I	F	Q	E	R	L	R	R	K	E	

FIG. 5